<center>**REMARKS**</center>

Claims 3-5, 7-9, 11-19, 21-25, 27-29, 31-33, 25 and 37-43 are pending in the application.

Claims 1-2, 6, 10, 20, 26, 30, 34 and 36 are cancelled above without prejudice.

Claims 3-5, 7, 9, 11-13, 16-17, 21-23, 27-29, 31-33 and 35 are amended above to clarify what it is that the applicant regards as the invention.

The application is amended to delete a redundant paragraph.

No new matter is added to the application by way of these specification amendments.

I.   THE ABSTRACT OBJECTION

The examiner objected to the application for lacking an Abstract on a separate page.

The Preliminary Amendment filed with the application on May 26, 2006 included a new Abstract at page 2 of the amendment. Therefore, the examiner is asked with withdraw the requirement for an Abstract because the application includes one on a separate page.

II.   THE CLAIM 17 OBJECTION

The examiner's objection to claim 17 is overcome by amending the claim as the examiner suggested in the Office Action.

III.   THE SECTION 112, 2nd PARAGRAPH REJECTION

The examiner rejected claims 1-39 for several reasons under 35 USC Section 112, 2nd paragraph. The examiner's rejections are overcome or traversed as set forth below.

- The examiner rejected claim 1 for using the allegedly indefinite term "proportion" in claim 1. The examiner also rejected claims 2-39 for depending upon claim 1 or for citing similar limitations.

    The Applicants have overcome the examiner's rejection by deleting claims 1 and 2 from the application and by replacing them with new claims 40 and 41 that do not include the objectionable language. Applicants have also deleted claims 6, 10, 20, 24, 26, 30, 34 and 36 from the application without prejudice, of which claims 20 and 30 are independent.

    Claims 3-19 now depend directly or indirectly from claim 40, and therefore no longer incorporate the deficiencies of claims 1 and 2.

<center>13</center>

Claims 20 and 30 have been deleted and replaced by independent claims 42 and 43 respectively, which are hardware and software claims having similar limitations to independent method claim 40. Claims 21-29 and 31-39 now depend directly or indirectly from claims 42 and 43 respectively.

Dependent claims 6, 10, 24, 26, 34 and 36 have been deleted and their corresponding limitations now appear in independent claims 40, 42 and 43.

- The examiner rejected claims 9, 25 and 35 for using the term "or otherwise. Claims 9, 25 and 35 have been amended to remove the objectionable language. The rejection of dependent Claims 10-15, 28-29 and 38-39 are overcome for the same reasons.

- The Examiner rejected claims 11, 27 and 37 for being unclear for using the expressions "variable" and "constant" to apparently refer to the same feature. In this regard the Oxford Dictionary of Computing, Oxford University Press, fourth edition, 1996, defines "variable" and "constant" as follows:

   Variable: A unit of storage that can be modified during program execution, usually by assignment or read operations. A variable is generally denoted by an identifier or by a name.

   Constant: A quantity or data item whose value doesn't change.

   In claims 11, 27 and 37, the expressions "variable" and "constant" are used to indicate that, during the rule refinement process, a variable (with appropriate mode declaration) can become replaced by a constant value – i.e. the variable thereupon stops being a variable and becomes a constant. For numeric values the constant is chosen from a range of potential values.

   Claims 11, 27 and 37 are amended above to clarify the claim language, i.e. "selecting a value having optimum accuracy": after "variable," these claims now continue "estimating a respective accuracy for each value and selecting for the variable one of the values which is optimum for accuracy in the range of values". It is believed that this amendment to claims 11, 27 and 37 overcomes the examiner's claim rejections.

- The examiner's rejection of claim 12 is overcome by amending the claim to specify that the second range of values has a narrow spacing relative to that of the first range; i.e. the limitations "relatively widely spaced" and "relatively narrowly spaced" have been deleted.

- The examiner's rejection of claim 13 is overcome by amending the claim to define a rule equivalent more clearly.
- The examiner's rejection of claim 16 is overcome by amending the claim to provide antecedent basis for "refinement encoding cost".
- The examiner rejected claims 28 and 38 because the claims do not define what "Iteq" means. In fact claim 14 defines "Iteq" which was originally thought to be adequate for later claims also. However, the definition is added to claims 28 and 38 above to overcome the examiner's rejection.

IV.   THE SECTION 101 REJECTION

The Examiner rejected claims 1-39 under 35 U.S.C. 101 for being directed to non-statutory subject matter. The pending independent claims are cancelled from the application above and replaced with new independent claims 40, 42 and 43. Claim 41 depends from claim 40, from which claims 3, 4, 7-9 and 11-19 also depend. Claims 21-23, 25 and 27-29 depend from claim 42 and claims 31-33, 35 and 37-39 depend from claim 43.

All three independent claims 40, 42 and 43 include the tangible result of "providing an alert or a report to a user regarding anomaly detection in the test data". All other claims depend from one or other of these independent claims, and so are likewise limited. The basis for this amendment appears in Applicants' specification at page 11 lines 20-22. As indicated by the Examiner in the official action, page 7 lines 2 and 3, presenting anomaly detection results to a user provides the required useful, concrete, and tangible result. Applicants trust that this overcomes the 35 U.S.C. 101 objection.

The Examiner also objects to the expression "any available relevant background knowledge data" in independent claims 1, 20 and 30 now deleted – although in fact "data" does not appear in those claims in conjunction with "any available relevant background knowledge": Applicants will treat the objection as being to "any available relevant background knowledge", and have amended it to "relevant background knowledge" in all new independent claims 40, 42 and 43.

The Examiner states that there is no description as to what this "relevant background data" is – Applicants will take this to be "relevant background knowledge". Applicants

respectfully traverse this statement, because Applicants' specification discloses numerous examples of relevant background knowledge – see e.g. the following:

page 11 lines 9-12:   "…definitions of background knowledge were generated in the form of concepts or functions relating to data attributes. One such function calculated a number of transactions handled by a specified cashier and having a discrepancy";

page 12 lines 22-26:   "…further information, i.e. concepts, facts of interest or functions that can be used to calculate values of interest e.g. facts about customers … accounts and a function ….. to calculate an average monthly bill … this further information is known as background knowledge…";

page 18 lines 23-24:   "A 'less than … are often used as part of the background knowledge";

page 24 line 6 to page 25 line 1: Background knowledge types are listed as "discrepancy between expected and actual cash in a till", "cash in that till for a particular transaction", "total number of transactions with a discrepancy in month/year, and "total number of transactions with a discrepancy greater than some bound in month/year";

page 28 lines 26-27:   "…concepts and facts of interest (background knowledge) expressed as Escher functions and constant values that may be extracted from training examples or specified by a user…";

page 31 line 25 to page 33 line 15: a small part of which is "The background knowledge in this case is that a discrepancy might be a useful thing to look at when constructing the rules: if so, it is necessary to find the size of the discrepancy that is sufficient to indicate fraud"; and

page 39 line 24 to page 41 line 15: This extract lists six items of background knowledge for use in assessing software vulnerability, see e.g. page 40 line 19: "This is an item of background knowledge which….".

      The Examiner further states that there is not "a restriction as to what this data is". Applicants have deleted the objected to "any" language from the claims. The Examiner also states that "the method may not necessarily consider the same data in order to develop the rules. Therefore, different results may be produced given the same data as inputs (not a repeatable result)". This is respectfully traversed. As indicated in Applicants' specification at page 11 lines 8-12, without access to experts who can supply background knowledge from their own experience, definitions of background knowledge can be generated using common sense: e.g. concepts or functions relating to data attributes such as a number of transactions handled by a

specified cashier and having a discrepancy between actual and expected cash. Consequently the invention can operate without expert background knowledge. If expert background knowledge is available, there will be additional material from which to develop rules from training data. In either event, the invention is repeatable: i.e. using the same training data and background knowledge, the same rule set will be developed and will produce the same anomaly results for reporting or alerting to a user. This is because logic is used in generating a rule set. Moreover, after a rule generalisation is processed to render it more specific, it is incorporated in the rule set if it classifies anomalies in the training data set adequately. More specific versions of rule generalisations are dropped if they don't satisfy this criterion.

Claims 3-19 and 41 depend from claim 40 and define patentable subject matter at least for the same reasons.

Claims 42 and 43 have similar limitations to claim 40 and are similarly patentable.

Claims 21-29 and 31-39 depend respectively from claims 42 and 43 and here again they define patentable subject matter at least for the same reasons.

Claim 30 has been deleted and replaced by claim 43, which now recites a computer readable hardware medium, which excludes an electrical or optical signal – see Applicants' specification at page 42 lines 17-19. Applicants trust that this embodiment now falls within one of the statutory categories of invention.

V.   THE ANTICIPATION REJECTION

Claims 1-39 are rejected under 35 U.S.C. 102(b) as allegedly being anticipated by US Patent No. 6,782,377 to Agarwal et al. (Agarwal). As has been said, claims 1, 2, 20 and 30 have been deleted and claims 3-19 now depend from claim 40. Claim 40 is not anticipated because Agarwal does not disclose the following limitations of claim 40:

(a)   a training data set expressed as digital data flagged to indicate presence and absence of data anomalies: it is in fact not clear how Agarwal determines whether a rare event is present or absent;

(b)   defining a rule generalisation based on logic of at least First-Order and processing to make it more specific by adding at least one of a condition, a variable, a constant, a unification of variables and a function based on the training data set and relevant background knowledge consisting of at least one of concepts, facts of interest and functions for calculating values of interest: in actual fact, Agarwal does not disclose even First-Order logic, let alone Higher-Order logic, but instead propositional logic (as confirmed below in more detail);

(c)   evaluating the more specific rule generalisation by applying it to the training data set to identify anomalies: instead Agarwal has a complicated technique for scoring P-Rule/N-Rule combinations – see C.9 L.50-53 and Appendix 3; or

(d)   incorporating the more specific rule generalisation in the rule set if it classifies anomalies in the training data set adequately in terms of covering positive anomaly examples.

Regarding Item (b) above, Agarwal's failure to disclose First-Order logic or Higher-Order logic is evidenced by C.1 L.49-56, which states:

> "The model for a given class C, is of the form R: D1 OR D2 . . . OR Dn-C, where each Di is called a disjunct, and is formed by conjunction of multiple conditions on attributes of the record, such as D1:(a1=P) AND (a2<10), where a1 and a2 are attributes of unordered and ordered types, respectively, and P is one of the specific values taken by a1."

Moreover, C.6 L.2-5 states:

> "Referring to FIG. 2, consider a binary classification problem, where a training dataset T and target class C form an input 202. For a rule R:A-*C (A being a conjunction of conditions formed by different attributes and their values),"

In the above two extracts from Agarwal, i.e., C.1 L..49-56 and C.6 L.2-5, the conditions are over a single attribute or variable and they are therefore propositional logic conditions. There are no relationships between or functions over multiple attributes (or variables), so the logic is neither Higher-Order logic nor even First-Order logic. In this regard the Oxford Dictionary of Computing, Oxford University Press, fourth edition, 1996 defines First-Order logic (or predicate calculus) as follows:

> "A fundamental notation for representing and reasoning with logical statements. It extends propositional calculus by introducing the quantifiers, and by allowing predicates and functions of any number of variables".

First-Order logic therefore cannot be propositional logic because it extends propositional calculus. Consequently, Agarwal does not disclose the invention as now claimed in claim 40, to which the 35 U.S.C. 102(b) rejection therefore does not apply. Claims 3-19 and 41 depend from claim 40, so Applicant's invention as now claimed therein is not disclosed by Agarwal either.

Regarding independent claim 20, as has been said this has been deleted and claims 21-29 now depend from independent claim 42 in which the objected to language has been amended. Claim 42 contains similar language to claim 40 and defines patentable subject matter for the

same reason recited above, which also applies to its dependent claims 21-29. Similar remarks apply to deleted independent claim 30 replaced by independent claim 43 with similar language to claim 40 and from which claims 31-39 now depend.

Regarding the objection to claim 20 relating to the expression "programmed to execute", replacement claim 42 now recites "... the computer apparatus incorporating computer software which programs it to execute..". Similar remarks apply to claims 22 and 26-29 now dependent claim 42. Applicants trust that this overcomes the objection.

Claims 3, 21 and 31 are not anticipated by Agarwal because they depend respectively from claims 40, 42 and 43 which are not anticipated for the reasons recited above.

The Examiner states that Agarwal anticipates developing the rule set using Higher-Order logic, referring to claims 4, 22 and 32: Applicants respectfully traverse this for reasons given above in relation to claim 40, i.e. Agarwal discloses propositional logic, not Higher-Order logic or even First-Order logic. Moreover, as before Claims 4, 22 and 32 are not anticipated by Agarwal because they depend respectively from claims 40, 42 and 43 which are likewise not anticipated.

The Examiner indicates that claims 5, 23 and 33 are anticipated by Agarwal. In particular, the Examiner states that "forming an alphabet having selector functions allowing properties of the training data set to be extracted" is disclosed in C.1, L.28 to C.2, L.14; C.3, L.10-59; C.6, L.1-56; C.7, L.28 to C.8 L.26; C.12-13, appendix 1-2, and "forming current rules from combinations of items in the alphabet such that type consistency and variable consistency is preserved" is disclosed in the abstract; C.3, L.10-59; C.6, L.11-56; C.13, and claim 1.

Applicants have deleted "at least one of the following:" from claims 5, 23 and 33, so that paragraph a) of these claims becomes "forming an alphabet having selector functions allowing properties of the training data set to be extracted, together with additional concepts, background knowledge constant values and logical AND and OR functions". The extracts from Agarwal cited by the Examiner do not mention "alphabet" or "selector functions" or "background knowledge": Agarwal discloses developing rules and models, not an alphabet. If the Examiner does not accept this, Applicants respectfully request the Examiner to specify the expressions in Agarwal which indicate or correspond to "alphabet", "selector functions" and "background knowledge", and the column(s) and individual line(s) (i.e. not column-length blocks) where they occur.

Claims 9, 25 and 35 are not anticipated by Agarwal because they depend respectively from claims 40, 42 and 43 which are not anticipated.

The Examiner states that Agarwal anticipates claims 11, 27 and 37, which relate to a variable in a rule in constant mode and numerical being at least partly evaluated by estimating an accuracy for each value in a range of values of the variable and selecting a value having optimum accuracy. In this regard the Examiner cites Agarwal C.4, L..16-49; C.6, L..1 to C.9, L..24, and refers to accuracy and support values and determining if a refined rule is accepted or if a current rule is best. It is respectfully submitted that Agarwal does not anticipate claim 11, 27 or 37, because two different concepts are being confused. Claims 11, 27 and 37 relate to deriving a value for a variable in a rule by selecting the value giving optimum accuracy. The Agarwal extracts cited by the Examiner relate to accuracy of a rule, not accuracy of a variable which is a component of a rule: (see *inter alia* Col. 6, line 27 of Agarwal) in one of the Examiner's cited extracts, "..rules that have high support and high accuracy". Claims 11, 27 and 37 are independently novel for at least this reason.

Similarly, claim 12 is not anticipated by Agarwal, because Agarwal does not disclose a range of values for a variable in a rule.

The Examiner states that Agarwal anticipates claim 13, which relates to filtering to remove rule duplicates and rule equivalents, and cites Agarwal: C.1, L.49 to C.2, L.14; C.6, L.1 to C.7, L.3. These cited extracts provide no disclosure whatsoever regarding removal of rule duplicates and rule equivalents or the like, and consequently the rejection of claim 13 is respectfully traversed.

The Examiner also states that Agarwal anticipates claim 14, which relates to filtering to remove unnecessary 'less than or equal to' ("Iteq") conditions, and cites Agarwal: C.1, L.49 to C.2, L.14; C.6, L.1 to C.7, L.3. Here again, the cited extracts provide no disclosure whatsoever regarding removing unnecessary Iteq conditions or the like, and consequently the objection is respectfully traversed. The Examiner observes that "After a rule is learned, the records where its antecedent is true are removed". However, here claim 14's deletion of redundant rules is being confused with Agarwal's removing (from the training dataset) data for which the rule is true, which is quite different (see e.g. C.6 L.21). Similar remarks apply to claim 15.

Remarks given above in relation to the novelty of claims 13 and 14 collectively apply equally to claims 28 and 38 which are apparatus and software equivalents of a combination of claims 13 and 14.

The examiner next takes the position that Agarwal anticipates claim 16, which relates to an encoding length restriction to avoid overfilling noisy data by rejecting a rule refinement having too high an encoding cost. The Examiner citing Agarwal C.3, L.45-59; C.7, L.29 to C.8, L.4 for supplying this teaching. Of these extracts, C.7 L.54-57 discloses a stopping criterion of the Description Length of the set of N-Rules not increasing beyond a pre-specified limit. However, this objection arises from Agarwal's length of an entire set of rules being confused with Applicants' length of an individual rule refinement, which might be a rule or part of a rule. Consequently the rejection is respectfully traversed as claim 16 is novel.

The Examiner further states that Agarwal anticipates claims 17, 29 and 39, which relate to stopping construction of a rule in response to fulfilment of at least one of three Agarwal stopping criteria, a) number of conditions in a rule, b) no negative examples are covered by a most significant rule, and c) no refinements produced eligible to enter the beam currently processed in a most recent refinement processing step. The examiner cited Agarwal extracts C.6, L.11 to C.7, L.18, C.7, L.28 to C.9, L.24 and. C.12-13, Appendices 1-2 for supplying this teaching. The examiner's rejection is respectfully traversed because Agarwal uses stopping criteria different to those of claims 17, 29 and 39. In this regard Agarwal C.9 L.5-24 discloses accepting R1 or stopping at R for P-Rules according to whether or not R1's EvaluationMetric is greater than that of R and enough training data examples are covered (MinSupFractionP). For N-Rules a rule grows until a false positive is covered and one less true positive is supported.

Regarding claim 18, the Examiner states that Agarwal anticipates adding the most significant rule to a list of derived rules and removing positive examples covered by it from the training data set, citing Agarwal C.1, L.49 to C.2, L.14; C.4, L.16-67; C.6, L.I to C.9, L.64; Fig. 2). This rejection is also respectfully traversed, because Agarwal uses different criteria for adding rules to a list of derived rules. In this connection, Applicants' specification at page 20 lines 18-22 defines a rule as 'significant' if its likelihood ratio statistic value is greater than a predefined threshold. Moreover, claim 18 depends from claim 17, which requires a most significant rule to cover no negative examples. In contradistinction, at C.6 L.23-24, Agarwal discloses P-Rules, i.e. rules added to the list of derived P-Rules. The P-Rules cover positive and

negative examples, and at C.6 L.41-44 Agarwal discloses N-Rules used to remove the false positives from the P-Rules. Consequently claim 17 is novel because the list of the P-Rules is specified in Agarwal before removing false positives, not afterwards as in claim 17. Dependent claim 18 is novel for this reason as well.

Finally, the Examiner states that Agarwal anticipates Claim 19, i.e. a) selecting rules which have not met rule construction stopping criteria, b) selecting a subset of refinements of the selected rules associated with accuracy estimate scores higher than those of other refinements of the selected rules, and c) iterating a rule refinement, filtering and evaluation procedure to identify any refined rule usable to test data. The Examiner cites in particular Agarwal C.4, L.16-67; C.6, L.1 to C.9, L.64; Fig. 2). However, it is respectfully submitted that this rejection arises from Agarwal's rule accuracy being confused with Applicants' accuracy of a refinement to a rule, which is quite different. Agarwal does not disclose accuracy of a refinement to a rule – see *inter alia* C.6 L.27 in one of the Examiner's cited extracts, "..rules that have high support and high accuracy". Claim 19 is independently novel for at least this reason.

CONCLUSION

The pending application and claims are believed to be in a patentable condition. Favorable reconsideration and allowance of all pending application claims is courteously solicited.

McDonnell Boehnen Hulbert & Berghoff LLP

Date: October 2, 2008

By: /A. Blair Hughes/
     A. Blair Hughes
     Reg. No. 32,901